Symbol Set Discriminability Metrics: Extending Discrimination Models for Size and Position Invariance

Albert J. Ahumada, Jr., Ph.D., Maite Trujillo San Martin, Ph.D., Jennifer Gille, Ph.D.
NASA Ames Research Center, Moffett Field, California

If flight display symbols are to be safely recognized by pilots, they need to be easily discriminated from each other. A study by Michael Zuschlag, DOT Volpe Center, assessed the recognizability of a proposed traffic symbol set. Predictions for the study results were generated by a standard image discrimination model. This model predicts that any difference whatever in the two images presented to it contributes to discriminability, while the observers appeared to categorize somewhat independently of size and position. An image discrimination model was developed that included both size compensation and position compensation. We applied this model to seven of the symbol pairs that lead to the most errors in the Volpe experiment. The predictions of experimental results by the model were improved. The model takes as input the luminance values for the pixels of two symbol images, the effective viewing distance, and gives as output the discriminability in just-noticeable-differences (d'), the size reduction of the larger symbol, and the x and y offset in pixels needed to minimize the discriminability.

INTRODUCTION

The goal of this project is to provide tools that can be used to evaluate the discriminability of symbols using extensions of visual discrimination models (Ahumada and Beard, 1998; Beard, Jones, Chacon, and Ahumada, 2005). Discriminability is only one component of the property Yey and Chandra (2004) call distinctiveness, the degree to which the symbol by itself can be identified. The visual discrimination models do not have a theory of feature learning or feature extraction or attention or memory effects. Bruner (1973) gives a good overview of these higher level processes that can affect symbol categorization. We are concerned here with symbol discriminability that only depends on low level visual processes.

Initially we planned to provide a model similar to that reported by Watson and Ahumada (2004; 2005). That model predicted the accuracy of letter identification in an acuity task as a function of optical distortions of the letters. We decided that in actual usage, the symbols would not be used with equal frequency and that actual performance correct was not as important as the possibility of potential errors. We decided to provide a tool that could be used to measure the discriminability of pairs of stimuli. All pairs in a set of potential symbols would need to be compared to ensure discriminability, but discriminability would not ensure accurate categorization. For example in the color domain, it is well known that many colors are discriminable from each other but that relatively few categories of colors can be accurately reported by naïve observers (Miller, 1956; Garner, 1962).

APPROACH

We began by looking at the data from the Volpe experiment, whose methodology and results were summarized by Zuschlag (2004).

"Methodology: The study is a descriptive psychophysical experiment. Ten pilots were recruited from a local airport. All had normal color vision and adequate visual acuity. The 19 symbols in the symbol set were presented one at a time on a bench-mounted aviation multifunction display (MFD) for 250 ms. The MFD was illuminated with approximately 94 kLx of light using a spotlight to simulate sun-shaft illumination. For each trial, each participant was shown a symbol in isolation and asked to select the perceived symbol from a matrix of 19 possible symbols presented on a laptop equipped with a touch screen. Error rates and reaction time were recorded.
Results: When viewed at a distance and angle approximating that found in a general aviation cockpit, most symbols were correctly recognized at least 92% of the time. The exception was symbols intended to indicate a selected state; these were correctly recognized as low as 83% of

the time. In the proposed symbol set, a selected state was indicated by outlining the symbol. The data suggest that this convention increases the likelihood that participants will confuse symbols indicating non-proximal traffic (represented by a hollow symbol) with symbols indicating proximal traffic (represented by a solid symbol)."

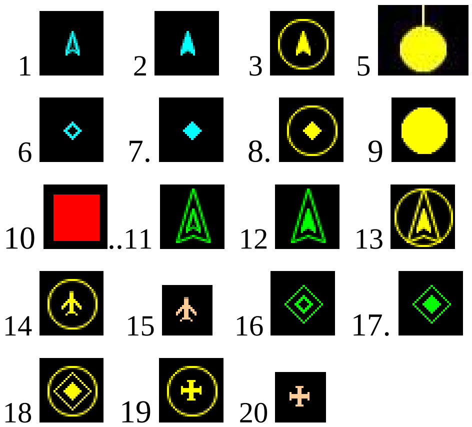

Figure 1. The Volpe experiment symbols.

Some of the results that appeared to us are the following:

1) There are significant order effects that are confounded with observer effects because the design was not balanced. Observer 5 made more than twice as many errors as anyone else, but he was the only one to begin with the farthest distance.

2) There are significant rotation effects. They only seemed to occur for the selected directional symbols. We will not be able to tell if these are perceptual or cognitive (categorical) without knowing what the actual images were (the rotation may have changed the outline). When Image 13 was in rotation 1 (we do not know what the positions mean), it was almost always called 13, but in other orientations it was often called 3 (see Figure 2).

Figure 2. Symbols 13 and 3.

3) The errors are not symmetric. A line version is called a filled version, but not the reverse. At the 88 in. distance, if wrong, Image 1 was most likely to be called 2, but Image 2 was most likely to be called 12 (see Figure 3), an indication of size invariance. At that distance, Image 11 was more likely to be called 12 than 11 (see Figure 3).

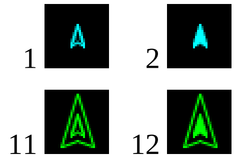

Figure 3. Image pairs 1 and 2 and 11 and 12.

The confusion of Image 13 being called 3 is also consistent with the filling-in principle and size invariance (see Figure 2).

Such principles imply that simple image difference models can not predict the actual pattern of responses. Another possible cause of response asymmetries is that the observer scanned through the responses sequentially, stopping when a match occurred without considering other possibilities.

4) Symbols 15 and 20 were mainly confused with each other, with a strong bias for responding that the plane (15) was present rather than the cross (20). At the 88 in. distance, they were actually more likely to say that the plane was present when the cross was. At that distance, the 2x2 confusion matrix has a d' of 0.6, while the image discrimination model predicts a d' of 1.1 (under various assumptions about the presentation conditions that we know are wrong). This result suggests the model can do a fair job of predicting the observed discriminability, since the observers vary in sensitivity by at least a factor of three.

Figure 4. Symbol pair 15 and 20.

From this analysis of the data from the Volpe experiment, we decided that a key problem with the existing discrimination models is that they do no compensate for pattern recognition transformational invariances that are naturally made by human observers. In the intended application of the symbols on a spatial map, the observers obviously must report that a symbol is the same when it is translated to a new position. Symbol size could be used as a cue, but, since the size would have to be anchored and is subject to strong context effects, we assume that symbols that are similar when one is magnified relative to the other will not be reliably discriminated. This transformation was suggested by the confusions between symbols 3 and 13 in the Volpe study despite the large discriminability predicted by the simple image discrimination model.

Size compensation is implemented by a frequency domain image shrinking algorithm (Watson, 1986). The current version only shrinks even sized images to even sized images, and the proportion lower range is an input parameter set to 0.5. This would have resulted in only 8 values of shrinkage for the 32 by 32 pixel symbols, so we pixel replicated the images by a factor of four so that 32 shrinkage levels were assessed. Higher resolution can be obtained by increased pixel replication of the images. The first pixel of the image must be the background level for the image because it is used to extend the smaller image so that all images are the same size before and after the size adjustment. Position compensation invariance is implemented by cross correlating the visual contrast images as a function of spatial offset by frequency domain filtering of one image by other. The pixel replication by a factor of four results in the position search being done to 0.25 pixel accuracy.

We have implemented Matlab routines to compute all the following steps:

1) Shrinking and padding of an image.

2) Conversion to visible contrast images.

5) Computation of the minimum visible difference position offset and the actual minimum visible difference there.

The code is available at (http://vision.arc.nasa.gov/personnel/al/code/index.htm).

RESULTS

Figure 5 shows some preliminary results from the model with size and position compensation. These calculations were done for full contrast images. Without size compensation, the model prediction of difference between symbols 3 and 13 was a d' of 4.8. With size compensation the predicted d' is lowered to 3.5. These d' values are slightly larger than those we previously reported without size and position compensation because the contrast sensitivity parameter was increased so that the best predicted detection performance for an observer would be 0 dBB (Watson, 2000).

The most interesting result was that the model now predicts that surrounding the plane and cross with a circle improves the discriminability (though not as much as observed). The plain image discrimination model with no masking predicts no effect of the circle. Adding masking to the model causes the model to predict that the circles will make the difference even more difficult to detect. Adding the translation invariance to the model allows the cross to be moved down to a better fitting position than is possible when the circle is present and stabilizes the position. The model thus pointed out that having the circle present allows one to see more clearly the asymmetric nature of the plane.

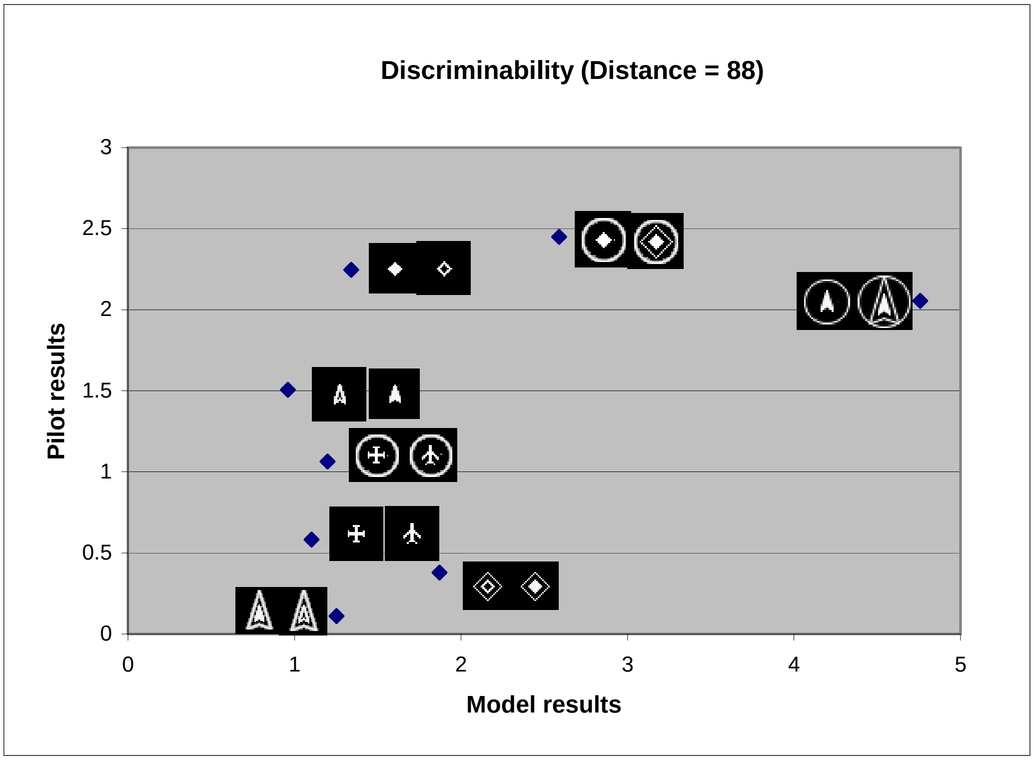

Figure 5. Predictions of the model with size and position compensation.

DISCUSSION

The implementation of the size compensation brought up two new issues that we had not considered before. One is that size compensation could be done separately in the x and y directions, so that the model would also predict confusions when a symbol is a taller or shorter version of another symbol. This feature would probably have helped the letter recognition predictions of Watson and Ahumada (2004). Another discovery is that when two images are different and about the same size there can be a slight advantage in having either of them made slightly smaller. So far, this effect has been small enough to be neglected.

There are two main issues that we have not dealt with. One is color and the other is orientation. In the Volpe experiment all the symbols were of a single color and there were essentially no confusions between differently colored symbols. For such symbol sets, the discriminability could be easily handled by adding color as a two more dimensions (essentially two more pixels). For symbol sets with multi-colored symbols, we would need to add two more images to each image in the manner of the detection model of Wuerger, Watson, and Ahumada (2002), together with the masking model of Ahumada and Krebs (2001)

As mentioned above, the Volpe study suggests that there were differences in the confusions as a function of orientation. If this result is not an artifact of rendering or of the fact that only one orientation was available in the response set, it would indicate that a simple model that extracted the orientation and the pattern from a “standard” orientation will not work. At this point the discriminability as a function of orientation can only be evaluated by brute force.

REFERENCES

Ahumada, A. J. Jr., Beard, B. L. (1998) A simple vision model for inhomogeneous image quality assessment, in SID Digest of Technical Papers, ed. J. Morreale, vol. 29, Paper 40.1, Santa Ana, CA.

Ahumada, A. J. Jr., and Krebs, W. K. (2001) Masking in color images, in B. E. Rogowitz and T. N. Pappas, Eds., Human Vision and Electronic Imaging VI, SPIE Proc. Vol. 4299, pp. 187-194.

Beard, B. L., Jones, K. M., Chacon, C., and Ahumada, A. J., Jr. (2005) Detection of blurred cracks: A step towards an empirical vision standard, Final Report for FAA Agreement DTFA-2045.

Bruner, J. S. (1973) Beyond the information given: Studies in the psychology of knowing. Oxford, UK: W. W. Norton.

Garner, W. R. (1962) Uncertainty and Structure as Psychological Concepts, New York: John Wiley and Sons, Inc.

Miller, G. (1956) The magical number seven, plus or minus two. Psychol. Rev. 63, 81-97.

Zuschlag, M (2004) Assessment of proposed traffic symbol set. FAA Human Factors Research and Engineering Division FY2004 Report, p. 44. https://www.hf.faa.gov/docs/508/docs/2004Report.pdf.

Watson, A. B. (1986) Ideal shrinking and expansion of discrete sequences, NASA Technical Memorandum 88202.

Watson, A. B., and Ahumada, A. J. (2004) Human optical image quality and the Spatial Standard Observer, OSA Fall Vision Meeting,

Watson, A. B., and Ahumada, A. J. (2005) Predicting acuity from aberrations with the spatial standard observer, Investigative Ophthalmology and Visual Science 46, E-Abstract 3614.

Wuerger, S., Watson, A. B., and Ahumada, A. J. Jr. (2002) Toward a standard observer for spatio-chromatic detection. SPIE Proc. Vol. 4662, 159-172.

Yeh, M., and Chandra, D. (2004) Issues in symbol design for electronic displays of navigation information. Proceedings of the 23rd DASC Conference, 24-28 October, Salt Lake City, UT.

ACKNOWLEDGEMENTS

This work was partly supported by the HMP Project of NASA's Airspace Systems Program. Dr. Trujillo is a National Research Council Associate and Dr. Gille was supported through the University Affiliated Research Center at NASA Ames Research Park.